A Fischer Inequality For The Second Immanant*

Robert Grone
Department of Mathematics
Auburn University, Alabama 36849

Russell Merris
Department of Mathematics and Computer Science
California State University
Hayward, California 94542

Dedicated to the memory of Emilie V. Haynsworth.

Submitted by Charles R. Johnson

ABSTRACT

Denote by $\mathscr{H}_n$ the cone of n-by-n positive semidefinite Hermitian matrices. Let d_2 be the generalized matrix function (or immanant) afforded by the symmetric group S_n and the irreducible degree $n-1$ character corresponding to the partition $(2,1,\ldots,1)$. Suppose $A \in \mathscr{H}_n$ is partitioned into blocks,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix}.$$

If $n \geqslant 4$, then $d_2(\mathrm{diag}(A_{11}, A_{22})) \geqslant d_2(A)$. This follows from the fact that d_2 is a Schur-concave function of the spectrum of A for $A \in \mathscr{C}_n$, where $\mathscr{C}_n$ consists of the matrices in $\mathscr{H}_n$ whose diagonal entries all equal 1.

*The work of both authors was supported in part by ONR contract No. 85-K-0335.

0024-3795/87/$3.50

INTRODUCTION AND RESULTS

Denote by $\mathcal{M}_n$ the space of n-by-n complex matrices. Let $\mathcal{H}_n \subset \mathcal{M}_n$ be the cone of positive semidefinite Hermitian matrices, and $\mathcal{C}_n \subset \mathcal{H}_n$ the compact convex subset consisting of those matrices all of whose main diagonal entries equal 1 (the *correlation matrices*).

Let G be a subgroup of the symmetric group, S_n, and let χ be an irreducible character of G. The *generalized matrix function* (GMF) afforded by G and χ is defined by

$$d(A) = \sum_{\sigma \in G} \chi(\sigma) \prod_{t=1}^{n} a_{t\sigma(t)}, \tag{1}$$

where $A = (a_{ij}) \in \mathcal{M}_n$. If $G = S_n$, then (following Littlewood) d is an *immanant*, and if $G = \{\mathrm{id}\}$, then d is the *Hadamard function*:

$$h(A) = \prod_{t=1}^{n} a_{tt}.$$

In case $\chi(\mathrm{id}) > 1$, it is convenient to define the *normalized* GMF, $\bar{d}(A) = d(A)/\chi(\mathrm{id})$.

We denote d_2 the immanant afforded by the character corresponding to the partition $(2,1,\ldots,1)$ of n. An explicit formula for χ is $\chi(\sigma) = \varepsilon(\sigma)(f(\sigma) - 1)$, where ε is the alternating or signum character and $f(\sigma)$ is the number of fixed points of σ. [Note that $\chi(\mathrm{id}) = n - 1$.] The immanant d_2 is referred to as the *second immanant* and satisfies ([4, §6.5] or [7])

$$d_2(A) = \sum_{t=1}^{n} a_{tt} \det A(t) - \det A, \tag{2}$$

where $A(t) \in \mathcal{M}_{n-1}$ is the principal submatrix of A obtained by deleting row and column t.

Recall that a function f of n nonnegative variables is (strictly) *Schur-concave* if $f(x) \le f(y)(f(x) < f(y))$ whenever x majorizes y and (and $y \ne x$). (See [6] for an introduction to majorization.)

THEOREM 1. *If $n \ge 4$, then $d_2(A)$ is a Schur-concave function of the spectrum of A on $\mathcal{C}_n$. Furthermore, d_2 is strictly Schur-concave on the matrices in $\mathcal{C}_n$ of rank at least $n-1$.*

THEOREM 2. *Suppose $A \in \mathcal{H}_n$ is partitioned as*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix},$$

where A_{11} and A_{22} are square. Let

$$\tilde{A} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}.$$

If $n \geq 4$, then $d_2(\tilde{A}) \geq d_2(A)$ with equality if and only if $A = \tilde{A}$, A has a zero row (and column), or rank $\tilde{A} < n - 1$.

Observe that Theorem 2 is an analog of the classical Fischer inequality for determinants. Not unexpectedly, it implies the following analog of the Hadamard determinant theorem.

COROLLARY. *If $A \in \mathcal{H}_n$, $n \geq 4$, then $\bar{d}_2(A) \leq h(A)$ with equality if and only if A is either diagonal or has a zero row (and column).*

PROOFS

Proof of Theorem 2. If $A = (a_{ij}) \in \mathcal{H}_n$ has a zero on the main diagonal, then A has a zero row and column. The proof is finished in this case. If every main-diagonal entry of A is positive, we can replace A with DAD, where $D = \operatorname{diag}(a_{11}^{-1/2}, a_{22}^{-1/2}, \ldots, a_{nn}^{-1/2})$. This has the effect of dividing both sides of the inequality by $h(A)$ and can be done without loss of generality. Since the spectrum of A strictly majorizes the spectrum of $\tilde{A}$ (unless $A = \tilde{A}$) [1], the inequality follows from Theorem 1. If $A \in \mathcal{H}_n$ is of rank $\geq n - 1$ and if A does not have a row of zeros, if follows from (2) and the Fischer inequality that $d_2(A) > 0$. Since rank $\tilde{A} \geq$ rank A, the case of equality is established. ∎

The corollary is a consequence of Theorem 2. Let $A_0 = A$ and

$$A_k = \left[\begin{array}{ccc|c} a_{11} & & 0 & \\ 0 & \ddots & & 0 \\ & & a_{kk} & \\ \hline & 0 & & B_k \end{array}\right],$$

$k = 1, 2, \dots, n$, where B_k is the principal $(n-k)$-by-$(n-k)$ submatrix of A obtained by deleting rows and columns $1, 2, \dots, k$. Then, for appropriate partitions, $A_{k+1} = \tilde{A}_k$.

[Indeed, $(n-1)h(A) = d_2(\operatorname{diag}(a_{11}, \dots, a_{nn})) \geq d_2(\tilde{A}) \geq d_2(A)$.] ■

Note that the corollary (and Theorem 2) fail when $n = 3$. Let

$$A_1 = \frac{1}{3}\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix} \quad \text{and} \quad A_2 = \frac{1}{3}\begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}.$$

then $\bar{d}_2(A_1) = \frac{28}{27}$, $\bar{d}_2(A_2) = \frac{26}{27}$, but $h(A_1) = h(A_2) = 1$. This example led the last two named authors of [3] to believe there was no direct comparison between h and $\bar{d}_2$ on $\mathscr{H}_n$.

Proof of Theorem 1: Using (2) and $a_{tt} = 1$, $1 \leq t \leq n$, we obtain that

$$\begin{aligned} d_2(A) &= \sum_{t=1}^{n} \det A(t) - \det A \\ &= E_{n-1}(\lambda) - E_n(\lambda), \end{aligned}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ are the eigenvalues of A, E_r is the rth elementary symmetric function, and $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$.

If rank $A \leq n-2$, then $d_2(A) = 0$. If rank $A = n-1$, then $d_2(A) = \lambda_1\lambda_2 \cdots \lambda_{n-1}$. Since this product is positive, strict Schur concavity is established in passing from rank $\leq n-2$ to rank $= n-1$. Among the rank-$(n-1)$ matrices, the product is strictly Schur-concave, and thus we are finished insofar as singular matrices are concerned.

(It may be worth interrupting the proof at this point to make the following observation in connection with the singular case. Since $\lambda_1 + \lambda_2 + \cdots + \lambda_{n-1} = n$, $d_2(A)$ will be maximized when $\lambda_1 = \lambda_2 = \cdots = \lambda_{n-1} = n/(n-1)$. But, as remarked in [2], this implies that $d_2(A) < e$, the base of the natural logarithms, for singular $A \in \mathscr{C}_n$. Hence there exists a bound for d_2 among the singular matrices in $\mathscr{C}_n$ which is independent of n.)

Returning to the proof of Theorem 1, we need only establish the strict Schur concavity among matrices of rank at least $n-1$. Since the only restriction on the eigenvalues of $A \in \mathscr{C}_n$ is that they majorize $(1, 1, \dots, 1)$, the proof reduces to the case $k = n-1$ in the following:

LEMMA. *Suppose n and k are positive integers, $n > k > (n+1)/2$. Let $R_{n,k} = \{(x_1, x_2, \ldots, x_n)\} : x_1 + x_2 + \cdots + x_n = n$ and $x_t \geq 0$, $1 \leq t \leq n$, with equality at most $n-k$ times$\}$. Define $F(x) = E_k(x) - E_{k+1}(x)$, $x \in R_{n,k}$. Then F is strictly Schur-concave on $R_{n,k}$.*

Since F is a symmetric function of the components of x, we may assume in our proof that $x_1 \geq x_2 \geq \cdots \geq x_n \geq 0$. Let e_t be the n-dimensional vector with a 1 in position t, and 0 in every other position. Define $u_t = e_{t+1} - e_t$, $1 \leq t < n$. By standard techniques (see, for example, [6, p. 55]), it suffices to show that the directional derivative of F in the u_t direction is positive whenever $x_t > x_{t+1}$. Now, the gradient of F is

$$\nabla F = \sum_{t=1}^{n} \left[E_{k-1}(x(t)) - E_k(x(t))\right] e_t,$$

where $x(t) = (x_1, \ldots, x_{t-1}, x_{t+1}, \ldots, x_n)$. Apart from the factor $\sqrt{2}$, the directional derivative in the u_t direction is

$$(x_t - x_{t+1})\left[E_{k-2}(x(t, t+1)) - E_{k-1}(x(t, t+1))\right],$$

where $x(t, t+1)$ is obtained from x by deleting x_t and x_{t+1}. Since we are assuming $x_t > x_{t+1}$, the proof will be complete upon showing

$$E_{k-2}(x(t, t+1)) > E_{k-1}(x(t, t+1)) \tag{3}$$

for $x \in R_{n,k}$ and $x_t > x_{t+1}$.

Now, it may happen that as many as $n-k$ components of $x(t, t+1)$ are zero. In the extreme case, only $k-2$ components of $x(t, t+1)$ are positive. But, in this case, the right-hand side of (3) is 0 while the left-hand side is positive. Thus, we may assume $x(t, t+1)$ contains at least $k-1$ positive terms. Denote by m ($\leq n-2$) the number of positive terms in $x(t, t+1)$. Let y be the m-triple of positive terms from $x(t, t+1)$. Then we need to show that $E_{k-2}(y) > E_{k-1}(y)$. But (see, for example, [5, p. 106])

$$\left[\frac{E_{k-2}(y)}{\binom{m}{k-2}}\right]^{1/(k-2)} \geq \left[\frac{E_{k-1}(y)}{\binom{m}{k-1}}\right]^{1/(k-1)}, \tag{4}$$

or

$$E_{k-2}(y) \geq \frac{\binom{m}{k-2}}{\binom{m}{k-1}^{(k-2)/(k-1)}} \left[E_{k-1}(y)\right]^{(k-2)/(k-1)}.$$

It remains to show that

$$\frac{\binom{m}{k-2}}{\binom{m}{k-1}^{(k-2)/(k-1)}} > \left[E_{k-1}(y)\right]^{1/(k-1)},$$

or

$$\frac{\binom{m}{k-2}}{\binom{m}{k-1}} > \left[\frac{E_{k-1}(y)}{\binom{m}{k-1}}\right]^{1/(k-1)}. \tag{5}$$

Several applications of (4) show that the right-hand side of (4)–(5) is dominated by

$$\frac{E_1(y)}{m} = \frac{n - x_t - x_{t+1}}{m} < \frac{n}{m}. \tag{6}$$

On the other hand,

$$\frac{\binom{m}{k-2}}{\binom{m}{k-1}} = \frac{k-1}{m-k+2}. \tag{7}$$

Comparing (6) and (7), the question boils down to whether $m(k-1) \geq n(m-k+2)$, or $(m+n)k \geq n(m+2)+m$. Since $k \geq (n+2)/2$, it is enough to show that $(m+n)(n+2) \geq 2n(m+2)+2m$, or $n \geq m+2$. As this is known to be correct, the proof is complete. ∎

REMARK. The lemma enables us to obtain stronger results than those stated. For example, in Theorem 1 we need only that A has diagonal entries all equal to 1 and nonnegative eigenvalues, *not* that A is Hermitian.

REFERENCES

1 K. Fan, Inequalities for eigenvalues of Hermitian matrices, *J. Res. Nat. Bur. Standards* 39: 131–139 (1954).

2 R. Grone, An inequality for the second immanant, *Linear and Multilinear Algebra* 18: 147–152 (1985).

3 C. R. Johnson, R. Merris, and S. Pierce, Inequalities involving immanants and diagonal products for H-matrices and positive definite matrices, *Portugal. Math.*, to appear.

4 D. E. Littlewood, *The Theory of Group Characters*, Oxford U. P., London, 1958.

5 M. Marcus and H. Minc, *A Survey of Matrix Theory and matrix Inequalities*, Allyn and Bacon, Boston, 1964.

6 A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic, New York, 1979.

7 R. Merris and H. Watkins, Inequalities and identities for generalized matrix functions, *Linear Algebra Appl.* 64: 223–242 (1985).

Received 18 July 1985; revised 30 January 1986